Quality Engineering

ISSN: 0898-2112 (Print) 1532-4222 (Online) Journal homepage: http://www.tandfonline.com/loi/lqen20

A Phase II nonparametric adaptive exponentially weighted moving average control chart

R. Zheng & S. Chakraborti

QUALITY ENGINEERING
2016, VOL. 28, NO. 4, 476–490
http://dx.doi.org/10.1080/08982112.2016.1183255

Taylor & Francis
Taylor & Francis Group

A Phase II nonparametric adaptive exponentially weighted moving average control chart

R. Zheng and S. Chakraborti

Department of Information Systems, Statistics and Management Science, University of Alabama, Tuscaloosa, Alabama

ABSTRACT
The in-control average run-length (ICARL) is often the metric used to design and implement a control chart in practice. To this end, the ICARL robustness of a control chart, that is, how well the chart maintains its advertised nominal ICARL value, under violations of the underlying assumptions, is crucial. Without the ICARL robustness, the shift detection property of the chart becomes questionable. In this article, first, the ICARL robustness of the well-known adaptive exponentially weighted moving average (AEWMA) chart of Capizzi and Masarotto (2003) is examined, in an extensive simulation study, with respect to the underlying assumption of normality. The ICARL profiles of the AEWMA chart are calculated for a range of distributions of various shapes, including light-tailed, heavy-tailed, symmetric, and skewed. Our results show that the AEWMA chart is quite sensitive to the normality (shape) assumption and may not maintain the nominal ICARL under non-normality. Motivated by this, a distribution-free (nonparametric) analog of the AEWMA chart (called the NPAEWMA chart), based on the Wilcoxon rank sum statistic, is proposed for Phase I applications when a Phase I reference sample is available. The NPAEWMA chart shows good ICARL-robustness against non-normality and shift detection capacity.

KEYWORDS
distribution-free; EWMA; ICARL; Phase I; robustness

Introduction

Traditional control charts in the statistical process control (SPC) literature include the Shewhart charts, the EWMA charts, and the CUSUM charts. The Shewhart chart is easy to implement in practice, but is less efficient in detecting smaller shifts, as it uses information from the current sample. By contrast, the EWMA and CUSUM charts use the information in a sequence of samples up to the point of comparison by combining the previous and present observations. These so-called averages-type charts are known to be effective for detecting small to moderate shifts with an assumed knowledge of the magnitude of the shift. In this article, we focus on the EWMA type charts as they are often preferred by the users. In practice, however, it may be difficult to design and use the EWMA chart effectively since the magnitude and the direction of the shift may be unknown and unpredictable so it's not clear what tuning (weight) parameters to use in the EWMA. In such situations, Capizzi and Masarotto (2003) (hereafter CM) considered an adaptive EWMA (AEWMA) chart, where the weighting parameter in the EWMA chart is calculated adaptively, over time, as each new data point becomes available. Thus, the AEWMA chart is able to adapt to changes in the process more efficiently and detect a variety of shifts, both small and large.

The idea of adaptively selecting the weighting parameter in the EWMA chart and improving chart performance is interesting and holds a lot of promise from a practical standpoint. Thus, over the last decade or so, a large amount of work has been done on the AEWMA chart and its various generalizations. In fact, a literature review found some 56 articles related to the AEWMA chart. Among these, Woodall and Mahmoud (2005) studied the inertial properties and evaluated the "signal resistance" of several control charts. They stated that "Likewise, the AEWMA procedure proposed by Capizzi and Masarotto (2003) has much better worst-case performance than the omnibus EWMA chart." Reynolds and Stoumbos (2006) compared different charts and chart combinations for monitoring the process mean and/or variance. They re-defined the "error" in the AEWMA by Capizzi and Masarotto (2003), and developed and evaluated AEWMA-type charts based

CONTACT S. Chakraborti schakrab@cba.ua.edu Department of Information Systems, Statistics and Management Science, University of Alabama, P.O. Box 870226, Tuscaloosa, AL 35487-0226.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lqen.

on squared deviations from target for monitoring process mean and/or variance. Shu (2008) also extended the idea of the AEWMA chart for monitoring process locations to the case of monitoring process dispersion. Mahmoud and Zahran (2010) proposed a multivariate AEWMA (MAEWMA) control chart to detect shifts in process mean vector. They made performance comparisons between the MAEWMA chart and the combined Shewhart-MEWMA chart in terms of the standard and worst-case average run length profiles and showed the effectiveness of their proposed MAEWMA chart. Simoes, Epprecht, and Costa (2010) compared the performance of the AEWMA, the combined EWMA-Shewhart scheme and the combined AEWMA-Shewhart scheme optimized for the same pair of shifts and concluded "First, there is no practical benefit in combining the AEWMA chart with a Shewhart chart. Second, the performances of the AEWMA chart and of the combined EWMA-Shewhart scheme are practically identical." More recently, Liu et al. (2013) proposed a sequential rank-based nonparametric adaptive EWMA (NAE) control chart for detecting the persistent shifts in the location parameter. Their NAE chart was claimed to be a self-starting procedure with no requirement of any prior knowledge of the underlying distributions, while CM's AEWMA chart was developed under the assumption of normality. Further, Liu, Tsung, and Zhang (2014) proposed a nonparametric adaptive CUSUM control chart based on the sequential rank as well. Saleh, Mahmoud, and Abdel-Salam (2013) pointed out that the AEWMA chart was proposed for the known parameter case, but in practice the process parameters are usually unknown and need to be estimated. It is well known that using estimates for the parameters degrades chart performance and hence Saleh, Mahmoud, and Abdel-Salam (2013) studied the performance of the AEWMA chart with estimated parameters. They showed the effect of different standard deviation estimators on the chart performance and recommended the use of the AEWMA chart over the standard EWMA chart especially when a small number of Phase I samples is available. In an interesting application setting, Tang et al. (2014) applied the AEWMA chart to detect the low-rate denial of service (LDoS) attacks on a network that reduce network service capability. They stated that "The NS2 simulations show that AEWMA method can detect LDoS attacks effectively and has a low false negative rate and a false positive rate. Based on DARPA99 datasets, experiment results show that AEWMA method is more efficient than EWMA method." Finally, Huang, Shu, and Su (2014) improved the computational method for estimating the run length performance of the AEWMA chart.

However, after a thorough review of these articles, it became apparent that none of these authors examined the ICARL robustness of the AEWMA chart to the distributional assumption of normality. By ICARL robustness we mean how well a chart maintains its advertised nominal in-control ARL value (such as 370 or 500) under violations of the assumption(s) (in this case normality) under which the chart was developed. It has been argued in the literature that without the ICARL robustness, the out-of-control shift detection property of a chart is somewhat meaningless. Note that we focus here on the violation of the normality assumption but violations of other assumptions (for other charts) can be of interest in the broader context of in-control robustness. Also, we consider here the ARL as the chart performance metric but other measures can also be considered.

It is clear that since in practice the underlying process distribution is most likely to be unknown or non-normal, without the in-control robustness, the practical value of any control chart may be highly diminished. This is because, for example, too many or too few false alarms (relative to the preset nominal value) can shatter the confidence of the user and ruin the efficacy of the chart. Hence, an in-control robustness study of any chart to the normality (or whatever distribution the chart is based on) assumption is essential before its implementation in practice. Liu et al. (2013) suspected the robustness of the AEWMA chart, stating "However, these control charts often assume that data come from some parametric distribution, most commonly the normal distribution. When the underlying process is unknown or not normal, these charts may not be appropriate", but they did not provide any evidence in support of their claim. In this article we study the in-control robustness of the AEWMA chart by examining its ICARL, as the ICARL is one of the most commonly used metric for control chart design and implementation in practice.

It should be noted that Borror, Montgomery, and Runger (1999) studied the ICARL robustness of the standard EWMA control chart to non-normality. They claimed that the standard EWMA chart can be designed so that it is robust to the normality assumption, so that the ICARL is reasonably close to the nominal for both skewed and heavy-tailed symmetric non-normal distributions. However, Human, Kritzinger, and Chakraborti (2011) also examined the robustness of the EWMA chart and suggested using caution against its overuse, particularly in situations where the shape of the underlying process distribution is not sufficiently known. But to the best of our knowledge, no one has examined the ICARL robustness of the AEWMA control chart to non-normality. In this article, we examine this important issue extensively, first using four standard distributions: the normal, the Laplace (the double exponential), the t and the uniform, and then six more distributions of different shapes within the g-and-k family of distributions. These six distributions include the (1) normal like, (2) symmetric heavy-tailed, (3) symmetric light-tailed, (4) slightly right-skewed, (5) slightly right-skewed and heavy-tailed, and (6) highly left-skewed distributions. Our results reveal that the AEWMA chart is highly sensitive to non-normality, so one has to be very careful while applying it in practice. We then adopt the basic idea behind the AEWMA chart and consider a nonparametric adaptation of the AEWMA chart, called the NPAEWMA chart, based on the Wilcoxon rank sum statistic. It is seen that the proposed distribution-free chart has attractive performance properties.

First, we start with a brief introduction to the AEWMA chart.

The AEWMA control chart

The basic idea of the adaptive EWMA (AEWMA) control chart, proposed by CM is to "adapt" the weights given to the past observations in a standard EWMA chart. More specifically, in an adaptive EWMA control chart the weighting parameter λ of the current observations changes along with every new observation coming in. In an AEWMA chart this constant weight λ is replaced by a suitable function of the current "error", which is the difference between the observed variable and the previous monitored value (the EWMA value at the previous time point). If the error value is small, the weight assigned to the current observation is small, and thus the chart can detect a small shift quickly, that is the chart behaves close to a standard EWMA chart. Otherwise, the current observation will be given a larger weight, and the chart will perform more like a Shewhart-type chart.

Assume that $y_1, y_2, \ldots, y_n$ are independent and identically distributed normal random variables. The AEWMA control chart is based on the statistics below:

$$\begin{aligned} x_t &= (1 - w(e_t))\, x_{t-1} + w(e_t)\, y_t \\ &= x_{t-1} + w(e_t)\left(y_t - x_{t-1}\right) \\ &= x_{t-1} + \phi(e_t), \end{aligned}$$

where $x_0 = 0$, $e_t = y_t - x_{t-1}$ is the "error", $w(e_t) = \phi(e_t)/e_t$, where $\phi(e_t)$ is the "score" function, which varies with the e_t. CM suggested three choices for the score function, among which the Huber score function is favored for its efficiency and simplicity. An AEWMA control chart with the Huber score function is the one that is most widely studied and was used by CM and also many researchers in the literature. The Huber score function is given by

$$\phi(e) = \begin{cases} e + (1-\lambda)k & \text{if } e < -k \\ \lambda e & \text{if } |e| \le k \\ e - (1-\lambda)k & \text{if } e > k \end{cases}.$$

The parameters λ and k are two of the three chart design parameters. The third chart design parameter is the control limit h (>0), that is chosen to guarantee a specified (nominal) in-control average run length value. The process is declared out of control when the monitored value x_t falls above h or below −h.

Recall that the key advantage of the AEWMA chart stems from the fact that it uses the flexible weighting scheme that helps it to adapt and adjust dynamically by "looking at" the successive differences between a "forecast" and the observation. However, as we noted earlier, the AEWMA chart is developed under the normality assumption, meaning that under non-normal process distributions the performance of the chart is not guaranteed. To study the in-control run length (RL) distribution of the AEWMA chart under non-normality, and hence its ICARL robustness, four well known distributions are used including the normal, the Laplace (the double exponential), the t, and the uniform. For further examination, six more distributions of different shapes are selected within the g-and-k distribution family (Hoaglin 1986; Haynes, MacGillivray, and Mengersen 1997). They are (1) normal like, (2) symmetric heavy-tailed, (3) symmetric light-tailed, (4) slightly right-skewed, (5) slightly right-skewed and heavy-tailed, and (6) highly left-skewed. We start with a description of the g-and-k family of distributions:

Table 1. Parameters for the six chosen g-and-k distributions.

Distribution Shape	gk1 normal	gk2 symmetric heavy tailed	gk3 symmetric light tailed	gk4 slightly skewed	gk5 slightly skewed heavy tailed	gk6 highly skewed
A	0	0	0	0	0	0
B	1	1	1	1	1	1
C	0.8	0.8	0.8	0.8	0.8	0.8
G	0	0	0	0.5	0.5	−2
K	0	0.5	−0.1	0	0.5	0

The g-and-k distribution family

The family of g-and-k distributions is defined by quantile function:

$$Q_x\left(u|A, B, G, K\right) = A + Bz_u\left(1 + C\frac{1 - e^{-Gz_u}}{1 + e^{-Gz_u}}\right) \times \left(1 + z_u^2\right)^K$$

where A and B >0 are the location and scale parameter, respectively, G is a measure of the skewness of the distribution, K>−0.5 is a measure of the kurtosis, z_u is the uth quantile of the standard normal distribution and C is a normalizing constant that produces a proper probability distribution. Approximately, $C \leq 0.83$ guarantees a completely proper distribution. Here we use $C = 0.8$ as researchers normally do. For more information about the properties of these distributions, see Haynes, MacGillivray, and Mengersen (1997) and Rayner (1999).

Table 1 displays the various parameters used in the study and Figure 1 shows the corresponding pdf's of the g-and-k distributions.

The ICARL robustness of the AEWMA chart

As for the design parameters (λ, k and h) in the AEWMA chart, we considered the two optimal

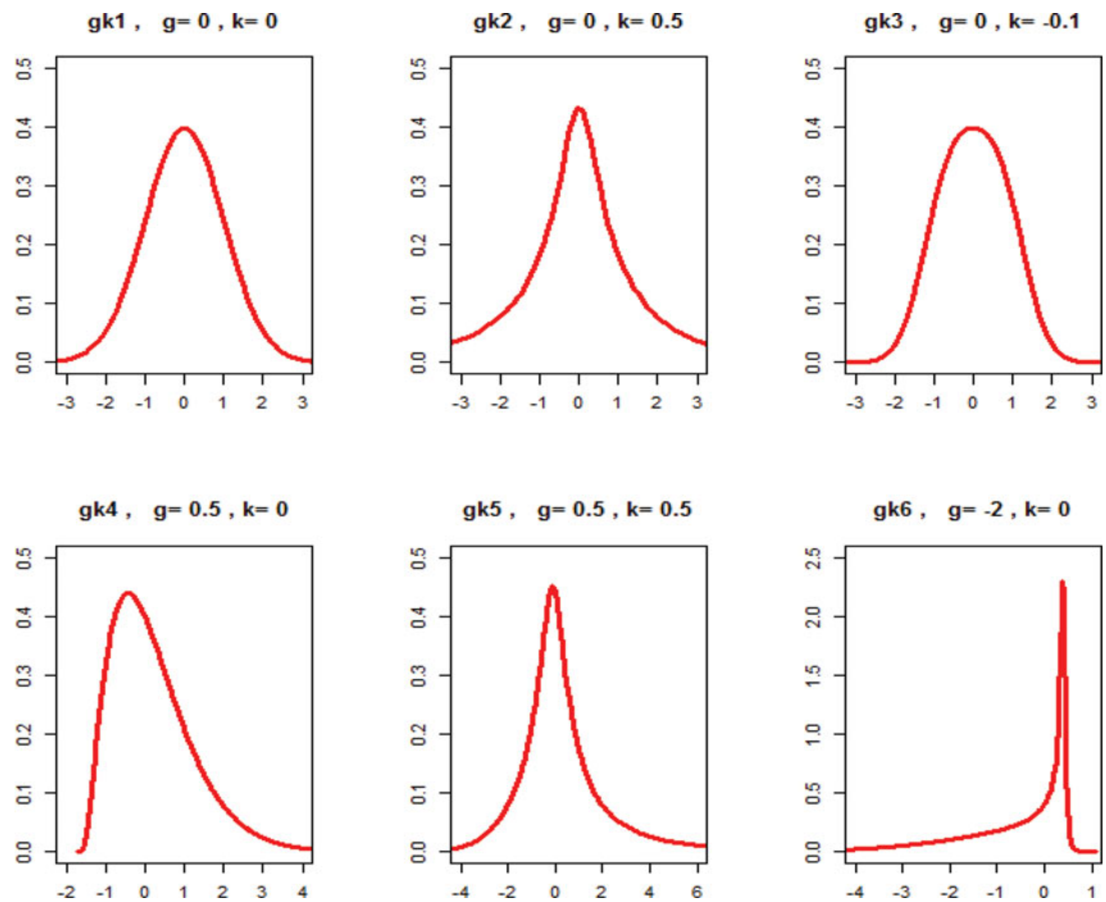

Figure 1. g-and-k distributions used in the simulations. From the left to the right and from the top to the bottom, the distributions are: normal like, symmetric heavy tailed, symmetric light tailed, slightly right skewed, slightly right skewed and heavy tailed, and highly left skewed, respectively.

Table 2. Optimal design parameters of two AEWMA charts for a nominal ICARL of 500.

Shift pairs		Optimal design parameters		
μ_1	μ_2	λ	k	H
0.5	4.0	0.0398	2.8990	0.4306
1.0	5.0	0.1354	3.2587	0.7931

Table 3. Optimal design parameters of two AEWMA charts for a nominal ICARL of 100.

Shift pairs		Optimal design parameters		
μ_1	μ_2	λ	k	h
0.5	4.0	0.0614	2.6306	0.3927
1.0	5.0	0.1913	3.2907	0.7688

combinations given by CM that produced an ICARL of 500 and 100, respectively, for the normal distribution. Tables 2 and 3 display these design parameter combinations used in our robustness study. Here a "shift pairs" denotes a pair of values of the shifted mean: a "small" and a "large", following CM.

It is well known that one can examine the zero-state or the steady-state performance of an EWMA control chart. In our study, as in CM, we investigate the robustness of the AEWMA chart in terms of the zero-state ICARL in a large number of Monte Carlo simulations. Tables 4a,b and 5a,b show the performance of the two AEWMA charts (for the two sets of chart parameters, for each nominal ICARL of 500 and 100, respectively, shown in Tables 2 and 3) for four well-known distributions: the normal, the Laplace (the double exponential), the t and the uniform, each centered and scaled if necessary so that the mean is 0 and standard deviation is 1, and six distributions of various shapes from the g-and-k distribution family: (1) normal like, (2) symmetric heavy-tailed, (3) symmetric light-tailed, (4) slightly right-skewed, (5) slightly right-skewed and heavy-tailed, and (6) highly left-skewed. All six g-and-k distributions also have their location parameters equal to 0, scale parameters equal to 1, and skewness/kurtosis as listed in Table 1. The results shown are based on 10,000 simulations.

The results in Tables 4 and 5 provide interesting insights into the ICARL robustness of the AEWMA chart. It is seen that the effect of the shape of the underlying distribution on the ICARL of the AEWMA chart is significant. As it might be expected, the attained ICARL values for the underlying normal or the normal like distribution are found to be very close to the nominal ICARL, when the underlying in-control distribution is not normal, the attained ICARL value can be vastly and significantly different from the nominal ICARL. For example, in Table 4a, for symmetric but heavier tailed t and Laplace distributions, the attained ICARL is much shorter than the nominal, leading to many false alarms whereas for the symmetric and lighter tailed uniform distribution, the attained ICARL is about three and a half times longer. For the g-and-k family, the same pattern is seen in Table 4b. For the symmetric heavy-tailed and highly left-skewed distribution, the attained ICARL values are significantly shorter than the nominal ICARL value or the symmetric light-tailed distribution, the attained ICARL values are remarkably longer than the nominal ICARL, say by more than 2000, which although results in a much lower false alarm rate than what is expected. This is not desirable either. The results in both of these tables show that, as designed, the AEWMA control chart performs well in the IC case of the normal distribution, but its performance is highly affected by non-normality, to the point that one cannot recommend this chart in practice

Table 4. The attained ICARL values of the AEWMA chart under various distributions; (a) normal, Laplace $(0,1/\sqrt{2})$, t_5, uniform $(-1/\sqrt{3},1/\sqrt{3})$, and (b) six different members of the g-and-k family with various shapes. The design parameters used are those as in CM for nominal ICARL of 500 under normality.

(a)

Shift pairs		Known distributions			
μ_1	μ_2	normal	Laplace	t5	Uniform
0.5	4.0	508.61	88.27	106.38	1937.67
1.0	5.0	504.22	130.28	145.09	961.74

(b)

Shift pairs		g-and-k family					
μ_1	μ_2	normal like	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
0.5	4.0	508.61	10.03	5337.44	70.83	10.93	23.15
1.0	5.0	504.22	11.68	2428.18	86.02	12.56	25.57

Table 5. The attained ICARL values of the AEWMA chart under various distributions; (a) normal, Laplace $(0,1/\sqrt{2})$, t_5, uniform $(-1/\sqrt{3},1/\sqrt{3})$, and (b) six different members of the g-and-k family with various shapes. The design parameters used are those as in CM for nominal ICARL of 100 under normality.

(a)

Shift pairs		Known distributions			
μ_1	μ_2	normal	Laplace	t5	Uniform
0.5	4.0	98.94	71.78	57.36	169.96
1.0	5.0	98.84	63.10	48.13	132.00

(b)

Shift pairs		g-and-k family					
μ_1	μ_2	normal like	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
0.5	4.0	99.72	7.86	296.45	37.42	8.66	15.16
1.0	5.0	98.76	9.09	254.72	43.80	9.91	16.55

without normality. This leads to the conclusion that, in general, the AEWMA chart is not ICARL robust.

Having observed the ICARL sensitivity of the AEWMA chart to the normality assumption, one can consider some possible remedies.

- Search for optimal design parameters which guarantee a nominal ICARL for a specified distribution. However, as described in CM, this solution will require a full knowledge of the underlying distribution. This is not practical.
- Consider pre-processing the data, such as applying transformation to normality and then applying the AEWMA control chart to monitor the process. We don't pursue this here.
- Adopt the basic idea of the adaptive EWMA chart but use a nonparametric/distribution-free statistic which does not require the normality assumption. This leads to a nonparametric analog of the AEWMA chart.

We pursue the third solution and propose a nonparametric version of the AEWMA chart, called the NPAEWMA chart, for the unknown parameter case, in the next section. The chart is based on the familiar Wilcoxon rank sum statistic.

The NPAEWMA control chart

Suppose $X = (x_1, x_2, \ldots x_m)$ and $Y = \left(y_1, y_2, \ldots y_n\right)$ denote two independent random samples from the distributions of two independent continuous variables. Wilcoxon (1945) proposed the rank sum test (WRS) based on the sum of the ranks (say V) of one of the samples, say the Y's, in the combined sample of the X's and the Y's. Let $a_{(i)} = 1$ if the ith smallest observation in the ordered combined sample is a Y and $a_{(i)} = 0$ otherwise. We can write $V = \sum_{i=1}^{n+m} \left(i \cdot a_{(i)}\right)$. Under the null hypothesis that the X's and the Y's are identically distributed, the expectation and the variance of V are (see, e.g., Gibbons and Chakraborti 2010): $E(V) = \frac{n(m+n+1)}{2}$ and $Var(V) = \frac{mn(m+n+1)}{12}$, respectively. It can be seen that if the distribution of the Y's is stochastically larger than that of X, then V will be large; otherwise V will be small.

To adopt this idea for a distribution-free control chart in the unknown parameter case, first, a reference sample from an in-control process needs to be obtained in Phase I. Then, at each point in time during the future monitoring of the process (Phase II), test samples (subgroups) are obtained and compared to the reference sample. A NPAEWMA control chart can be constructed as follows.

- Collect a reference sample of size m: $X = (x_1, x_2, \ldots, x_m)$, from an in-control process. Note that the available Phase I data is in a single sample of size m.
- Collect subgroups, each of size n, at time t, from the monitored process. Let the subgroup be denoted by $Y_t = \left(y_1, y_2, \ldots, y_n\right)$.
- Compute the Wilcoxon rank sum statistic V_t between the Y and the X sample at time t. Let $V'_t = \frac{V_t - E(V)}{\sqrt{var(V)}}$ be the standardized V_t.
- Adapt the rank sum statistic in the AEWMA framework and calculate the NPAEWMA monitoring statistic $T_t = (1 - w(e_t))T_{t-1} + w(e_t)V'_t$, with $T_0 = 0$ and $w(e_t)$ as for the AEWMA chart.
- An out-of-control signal is given when T_t falls on or outside the control limits, that is, on or outside of $\pm\, h$.

Table 6. The attained ICARL values of the NPAEWMA chart for six distributions. The design parameters used are those for the AEWMA control charts for the same shift pairs and the nominal ICARL of 500. Number of simulations = 10,000.

Shift pairs		Underlying distributions					
μ_1	μ_2	normal	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
					m = 100, n = 25		
0.5	4.0	466.87	462.31	480.89	462.79	464.45	483.76
1.0	5.0	446.97	453.86	445.55	443.91	461.98	429.74
					m = 500, n = 5		
0.5	4.0	800.06	802.06	798.20	818.19	794.68	796.24
1.0	5.0	506.31	508.58	506.48	503.18	506.16	508.39

- The design parameters (λ, k, h) used for the NPAEWMA chart, for the first time, are taken to be the same as those for the AEWMA chart of CM.

Note that a Shewhart-type distribution-free chart, based on V_t was considered by Chakraborti and Van de Wiel (2008).

The ICARL robustness of the NPAEWMA control chart

Under the IC situation, when the distributions of X and Y_t are identical, it is well known (see, for example, Gibbons and Chakraborti 2010) that the distribution of V'_t is approximately standard normal, when m and n are sufficiently large. The Wilcoxon rank sum test based on V_t (or V'_t) is distribution-free, that is, no matter what the common continuous distribution of X and Y_t is, the in-control distribution of V_t is only a function of m and n. By definition, a nonparametric control chart based on V'_t is distribution-free and so it should be ICARL robust. However, since the distribution of V'_t is standard normal for large m and n, it is of interest to examine the (finite sample) ICARL performance of the NPAEWMA chart as a function of m and n. To this end, two combinations of m and n values are selected for each of the six g-and-k different distributions considered earlier. Note that in the NPAEWMA charts we first use the same optimal charting parameters (λ, k, and h) of the AEWMA chart and thus consider a direct nonparametric analog. The results are shown in Tables 6 and 7.

From Tables 6 and 7, and comparing with the results in Tables 4b and 5b, it is seen that the NPAEWMA chart is much more robust to non-normality when compared with the AEWMA chart. For example, in Table 6, when m = 500 and n = 5, the NPAEWMA control chart, for shift pair 1 and 5, has ICARL values ranging from 503.18–508.58, for a nominal ICARL value of 500, for all six distributions. Whereas for m = 100, n = 25, the NPAEWMA control chart for shift pair 0.5 and 4 has ICARL values ranging from 462.31–483.76. It is clear that the differences between the attained and the nominal ICARL values, for different shapes of distributions are significantly smaller for the NPAEWMA chart than for the AEWMA control chart. By comparison, for the AEWMA chart, these differences range from tens to thousands. For example, from Table 7, when m = 500 and n = 5, for the NPAEWMA control chart, for shift pair 1 and 5, the attained ICARL values range from 95.39–97.01 when the nominal ICARL is 100, but from Table 5, for the same situation with the AEWMA control chart, the ICARL values range from 9.09–254.72. In summary, a really wide range of ICARL values with some very small and some very large values point to the weakness of the AEWMA chart and point to the stability and the ICARL superiority of the nonparametric analog, the NPAEWMA chart.

Table 7. The attained ICARL values of the NPAEWMA chart for six distributions. The design parameters are those for the AEWMA control charts for the same shift pairs and the nominal ICARL of 100. Number of simulations = 10,000.

Shift pairs		Underlying distributions					
μ_1	μ_2	normal	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
					m = 100, n = 25		
0.5	4.0	71.80	71.75	71.28	73.96	71.57	71.61
1.0	5.0	79.00	79.07	78.69	79.52	78.73	81.30
					m = 500, n = 5		
0.5	4.0	108.45	104.71	105.85	104.58	104.83	106.22
1.0	5.0	97.01	95.39	95.78	95.55	95.97	95.28

Table 8. Results of the Shapiro-Wilk test for normality about the standardized WRS statistic V'_t.

m	n	$\frac{m}{m+n}$	number of rejections among 500 repetitions
100	5	0.952	464
100	15	0.870	39
100	25	0.800	24
300	5	0.984	453
300	15	0.952	33
300	25	0.923	28
500	5	0.990	439
500	15	0.971	44
500	25	0.952	22

However, it may be noted that despite the vast superiority on the basis of the ICARL robustness to non-normality, one important issue related to the IC performance of the NPAEWMA control chart is that all the attained ICARL values (in Tables 6 and 7) are somewhat less (more) than the nominal ICARL of 500(100). There are two reasonable explanations.

1) The AEWMA control charts are designed for process monitoring with known parameters, while the distribution-free NPAEWMA chart is used for process monitoring with unknown parameters, where a reference sample of size m is assumed to be available from an in-control process. Naturally, a large number of reference sample observations are needed in order to maintain the ICARL at the nominal level. In the following sections of this article, simulated corrected control limits are used to guarantee the actual ICARL reaches the nominal ICARL.
2) There may be questions about the in-control distribution of the statistic V'_t, which, in theory, is standard normal, for large m and n values. This also points to needing a sufficient amount of reference data.

To explore the second point further, Table 8 presents the results of the Shapiro-Wilk test for normality applied to 5,000 values of the V'_t statistic, calculated for certain chosen values of m and n with both samples from the standard normal distribution. The steps are as follows.

(1) Choose m and n.
(2) Randomly generate m reference observations X, and n test observations Y, both from the standard normal distribution.
(3) Calculate the standardized WRS statistic for the Y sample, and denoted it as the V'_t.
(4) Repeat steps 2 and 3 until 5000 V'_t values are obtained.
(5) Perform the Shapiro-Wilk test for normality and record the p-value for this one single test.
(6) Repeat steps 2–5 until we have 500 p-values from the Shapiro-Wilk test.
(7) Count the number of p-values smaller than or equal to 0.05 among the 500.
(8) Repeat steps 1–7 for different m and n values.

Table 8 shows the results for m = 100, 300, 500, and n = 5, 15, and 25, respectively. As it might be expected, it is seen that when n = 5, the normal approximation of V'_t is not as good as when n = 25, since in this case the number of rejections varied around 25. Thus, n = 25 should be preferred by practitioners. However, from the point of view of a quality practitioner, a subgroup size of n = 5 is more attractive and has been widely used. The performance of the proposed chart when n = 5, will be examined in the next section.

The performance of the NPAEWMA control chart

Having established the ICARL robustness of the NPAEWMA chart, in this section, the out-of-control performance of the NPAEWMA chart is evaluated here. We also investigate the effect of the reference sample size m and the test subgroup size n on the chart performance. To this end, note that as discussed in the previous section, using the design parameters intended for the AEWMA chart on the NPAEWMA chart tend to produce ICARL values closer to the nominal value but the closeness can be improved. In order to make the OOC comparisons fair, the charts under comparison should have roughly the same ICARL. One suggestion by many authors, e.g., Mahmoud and Maravelakis (2010) and Jones (2002), to this end is to adjust the control limits empirically. The same is done here and these adjusted control limits are referred to as the "corrected control limits." Thus, the control limits in Table 2 and Table 3 are adjusted for producing an ICARL of 500 or 100, respectively, corresponding to m = 100, n = 25 and m = 500, n = 5 for the NPAEWMA chart. Table 9 shows the corrected control limits h*. The design parameters λ and k in the AEWMA remain the same.

The corrected control limits h* produce an ICARL close to the nominal level for all six distributions of different shapes. Note that the corrected control limits for the NPAEWMA chart (the h*) are somewhat close to the control limits (the h) for the AEWMA control chart

Table 9. The corrected control limits h^* required to produce an ICARL of 500 for six distributions. The reference sample size and subgroup size are m = 100, n = 25 and m = 500, n = 5. The values of k and λ are those used for the AEWMA control chart.

Shift pairs		ICARL = 500				
μ_1	μ_2	λ	k	h	corrected control limits h^*	
0.5	4.0	0.0398	2.8990	0.4306	m = 100, n = 25	0.4353
					m = 500, n = 5	0.4000
1.0	5.0	0.1354	3.2587	0.7931	m = 100, n = 25	0.8036
					m = 500, n = 5	0.7931

Table 10. The corrected control limits h^* required to produce an ICARL of 100 for six distributions. The reference sample size and subgroup size are m = 100, n = 25 and m = 500, n = 5. The values of k and λ used are those for the AEWMA control chart.

Shift pairs		ICARL = 100				
μ_1	μ_2	λ	k	h	corrected control limits h^*	
0.5	4.0	0.0614	2.6306	0.3927	m = 100, n = 25	0.4253
					m = 500, n = 5	0.3896
1.0	5.0	0.1913	3.2907	0.7688	m = 100, n = 25	0.7946
					m = 500, n = 5	0.7747

given by CM. In general, the h^* values are larger than h, except for the charts for which the shift pair is 0.5 and 4 and m = 500, n = 5, where the h^* is smaller than h, meaning that in these two cases, the control limits h are narrowed to meet the nominal level. For the case where the shift pair is 1 and 5 and m = 500, n = 5, the corrected h^* is the same as the control limit h. For the rest of the cases, we expand the control limits a little bit to meet the nominal level. The search of the h^* is done by computers by first setting a range around h and then simulating the monitoring process against all the h values within the chosen range. The h^* is the control limit that has an ICARL close to 500/100 for all distributions. Results show that only small modifications to h are needed.

Tables 11 and 12 display the out-of-control average run length (OOCARL) values of the NPAEWMA control chart for the six distributions, when the corrected control limits, obtained from Tables 9 and 10, are used. Note that the shifts are expressed in the units of $1/\sqrt{n}$

Table 11. The OOCARL values of the NPAEWMA chart for the six distributions. The nominal ICARL equals 500. Each ARL is calculated from 10,000 simulations of run lengths. Two combinations of m and n are used: m = 100, n = 25 and m = 500, n = 5.

Shift pairs		Shift in location	Underlying distributions					
μ_1	μ_2	δ	normal	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
		in unit of $1/\sqrt{25}$			m = 100, n = 25			
0.5	4.0	0.00	523.52	505.56	520.98	497.62	530.92	515.26
		0.25	454.78	482.21	461.18	502.57	514.41	64.01
		0.50	315.40	433.67	315.16	394.04	459.78	14.00
		1.00	93.59	222.18	80.96	128.19	245.91	5.32
		1.50	19.65	90.41	17.90	21.57	107.84	3.01
		2.00	6.93	25.00	6.21	7.02	37.62	1.84
		2.50	4.13	11.64	3.76	3.81	11.69	1.19
		3.00	2.57	5.85	2.24	2.31	6.01	0.76
1.0	5.0	0.00	508.82	521.12	482.31	502.99	497.88	479.95
		0.25	469.84	478.07	430.09	491.99	483.97	64.22
		0.50	350.00	401.35	316.22	407.44	434.13	11.65
		1.00	101.75	219.60	88.94	134.04	263.95	4.04
		1.50	16.61	77.63	13.17	21.28	108.85	2.53
		2.00	5.11	30.37	4.51	5.76	34.69	1.74
		2.50	3.03	7.99	2.76	2.71	10.94	1.26
		3.00	2.08	4.34	1.90	1.86	4.93	0.94
		in unit of $1/\sqrt{5}$			m = 500, n = 5			
0.5	4.0	0.00	511.53	503.79	506.97	518.84	497.83	509.62
		0.25	159.45	259.36	145.03	148.99	261.71	22.47
		0.50	38.15	74.52	35.58	34.03	71.77	11.50
		1.00	12.95	20.86	12.35	11.94	20.00	6.19
		1.50	7.55	12.03	7.11	7.20	11.61	4.22
		2.00	5.07	8.57	4.71	5.10	8.24	3.13
		2.50	3.48	6.52	3.16	3.81	6.40	2.39
		3.00	2.32	5.21	2.05	2.83	5.27	1.84
1.0	5.0	0.00	500.97	505.88	507.21	505.26	506.34	507.10
		0.25	198.33	289.33	181.37	202.28	312.05	21.50
		0.50	45.78	98.67	41.68	42.92	96.54	9.88
		1.00	10.58	20.23	9.80	9.53	19.48	5.06
		1.50	5.45	9.48	5.09	4.94	9.10	3.46
		2.00	3.56	6.20	3.35	3.31	5.76	2.59
		2.50	2.63	4.47	2.48	2.49	4.22	2.05
		3.00	2.06	3.57	1.93	2.05	3.37	1.66

Table 12. The OOCARL values of the NPAEWMA chart for the six distributions. The nominal ICARL equals 100. Each ARL is calculated from 10,000 simulations of run lengths. Two combinations of m and n are used: m = 100, n = 25 and m = 500, n = 5.

Shift pairs		Shift in location	Underlying distributions					
μ_1	μ_2	δ	normal	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
		in unit of $1/\sqrt{25}$				m = 100, n = 25		
0.5	4.0	0.00	97.68	103.21	98.39	100.97	100.18	98.72
		0.25	90.42	93.98	86.98	94.98	96.79	25.20
		0.50	68.79	83.15	63.72	71.85	85.08	8.20
		1.00	26.10	50.34	22.65	25.87	50.10	3.26
		1.50	8.49	22.08	7.86	8.31	24.70	1.85
		2.00	4.04	9.37	3.71	3.67	9.48	1.17
		2.50	2.45	5.60	2.12	2.19	5.86	0.73
		3.00	1.45	3.46	1.27	1.25	3.48	0.48
1.0	5.0	0.00	97.42	96.30	98.69	96.93	97.62	97.22
		0.25	87.48	92.93	88.09	90.80	94.78	23.59
		0.50	67.41	82.73	64.60	71.84	84.10	7.04
		1.00	25.44	47.96	21.69	26.16	50.48	2.69
		1.50	7.28	20.48	6.56	7.03	21.67	1.67
		2.00	3.28	9.37	2.88	2.81	9.40	1.18
		2.50	1.94	4.44	1.78	1.68	4.56	0.85
		3.00	1.34	2.72	1.22	1.14	2.79	0.64
		in unit of $1/\sqrt{5}$				m = 500, n = 5		
0.5	4.0	0.00	102.60	101.20	103.69	101.19	104.61	102.46
		0.25	53.63	70.77	49.95	52.86	71.72	12.75
		0.50	21.24	34.00	19.31	19.47	33.05	7.01
		1.00	7.81	12.30	7.26	7.18	12.08	3.73
		1.50	4.40	7.16	4.11	4.18	7.00	2.55
		2.00	2.84	5.02	2.64	2.85	4.83	1.87
		2.50	1.86	3.66	1.69	1.94	3.66	1.39
		3.00	1.16	2.79	1.01	1.22	2.85	1.07
1.0	5.0	0.00	98.48	98.97	101.13	99.53	100.71	100.77
		0.25	57.38	74.05	53.26	58.53	73.87	12.51
		0.50	22.03	36.67	20.27	20.81	36.78	6.44
		1.00	6.63	11.67	6.28	6.04	11.48	3.40
		1.50	3.49	6.11	3.26	3.17	5.73	2.33
		2.00	2.23	3.89	2.12	2.04	3.70	1.76
		2.50	1.61	2.86	1.50	1.48	2.67	1.37
		3.00	1.21	2.24	1.13	1.14	2.09	1.13

where n is subgroup size, because the standard deviation for the subgroup mean is $B/\sqrt{n}$, with B = 1, which is the scale parameter in the g-and-k distribution. Thus, for example, a shift of $\delta = 0.25$ in tables below means a shift of $0.25 \times 1/\sqrt{25} = 0.05$ for n = 25 or $0.25 \times 1/\sqrt{5} = 0.11$ for n = 5, respectively, in the location.

In practice, the combination of m = 500 and n = 5 is recommended for the NPAEWMA chart when parameters are unknown. This is in line with the recommendation for the normal theory control charts as recent research has shown (see, e.g., Saleh et al. (2015), Epprecht, Loureiro, and Chakraborti (2015)). However, it should be noted that in this article m denotes the number of Phase I observations. Generally speaking, it is seen that for a heavier-tailed distribution relatively more data are required to detect a shift whereas much less data are needed for a highly skewed distribution. This seems to hold in all the situations we examined. When m = 100, n = 25, the NPAEWMA chart for the shift pair 1 and 5 performs better than the one for the shift pair 0.5 and 4 for both small and large magnitudes of shifts. When m = 500, n = 5, the chart for shift pair 0.5 and 4 is better for smaller shifts whereas the chart for shift pair 1 and 5 is better for larger shifts.

Based on our results, for a nominal ICARL = 500, we recommend m = 500, n = 5, $\lambda = 0.1354$, k = 3.2587, $h^* = 0.7931$ for the NPAEWMA chart for its robust ICARL performance and better OOC performance. However, although m = 500 is an acceptable Phase I sample size in the most recent literature based on the analysis of effects of parameter estimation on control charts, it is still a relatively large number. Practitioners may not have the time to wait until this much data are gathered, and in the meantime, the production process may have a higher risk of going out of control. Thus, we also investigate the ICARL performance of the NPAEWMA chart for both sets of parameters (shift pair 0.5 and 4 and 1 and 5, respectively) using perhaps a more reasonable m = 100 and n = 5 and using corrected control limits, when the nominal ICARL is 500. Table 13 displays the attained ICARL values in this case.

Table 13. Attained ICARL values for a reference sample size m = 100 and n = 5. Each ICARL is calculated from 10,000 simulations of run lengths using the corrected control limits.

nominal ICARL = 500, m = 100, n = 5, with corrected control limits h*							
Shift pairs		Underlying distributions					
μ_1	μ_2	normal	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
0.5	4.0	λ = 0.0398, k = 2.8990, h* = 0.4000					
		388.65	386.72	378.77	382.72	379.15	396.34
1.0	5.0	λ = 0.1354, k = 3.2587, h* = 0.7931					
		471.16	480.12	473.20	473.36	487.02	471.71

Table 14. The attained ICARL values of the NPAEWMA chart for six different shapes of underlying distributions for m = 100 and n = 1. The design parameters are those for the AEWMA control charts for the same shift pairs and the nominal ICARL is 500/100. Number of simulations = 10,000.

Shift pairs		Underlying distributions					
μ_1	μ_2	normal	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
		m = 100, n = 1, ICARL = 500					
0.5	4	1237.57	1241.01	1215.06	1254.42	1265.15	1247.88
1	5	829.43	831.52	831.59	817.39	842.28	838.76
		m = 100, n = 1, ICARL = 100					
0.5	4	137.96	137.1	137.61	137.58	137.57	136.98
1	5	122.35	121.71	121.35	123.63	125.21	122.41

Table 13 shows, for situations where not enough reference sample data are available, the attained ICARL values are less than the nominal value, which will produce a higher false alarm rate. However, the NPAEWMA chart for shift pair 1 and 5 that is with (λ, k, h*) = (0.1354, 3.2587, 0.7931) is reasonably robust and may be used in practice.

Performance of NPAEWMA for individual observation and some comparisons with other charts

So far we have considered two Phase II subgroup sizes, n = 5 and n = 25. However, sometimes smaller subgroup sizes may be preferred. For example, in some applications it is common to monitor individual data, that is, a sample of size n = 1, since monitoring individual observation may be natural and perhaps cost efficient. Thus, it is of interest to examine the performance of the proposed NPAEWMA control chart for individual Phase II data. Table 14 shows the robustness of the NPAEWMA control chart for six different distributions. The chosen parameters (λ, k and h) are the same as we used for n = 5 and n = 25 cases, with m = 100, n = 1 and nominal ICARL = 500/100.

The conclusions from Table 14 are consistent with those from Tables 6 and 7. The proposed NPAEWMA chart is robust against non-normality when monitoring individual observations. But the actual ICARL is not the same as the nominal ICARL. This is most likely due to the same reason as in Section 6 where we simulated the ICARL for the n = 5 and n = 25 cases. Thus, Table 15 provides the corrected control limits, for n = 1, for the NPAEWMA chart to reach the nominal ICARL. Table 16 shows the OOC performance when monitoring individual observations.

Liu et al. (2013) considered a sequential rank-based nonparametric adaptive control chart. Their chart uses a change-point formulation and monitors the individual data. In their article, they made performance comparisons with the modified EWMA chart with "standardized ranks" proposed by Hackl and Ledolter (1991), a CUSUM procedure based on "sequential ranks" proposed by McDonald (1990), and a nonparametric change-point chart proposed by Hawkins and Deng (2010). These three charts are denoted as HLE, McC, and HDC, respectively. Note that our

Table 15. The corrected control limits h* for the NPAEWMA chart required to produce an ICARL of 500/100 for six distributions. The reference sample size is m = 100 and the subgroup size is n = 1. The values of k and λ are those used for the AEWMA control chart.

Shift pairs		m = 100, n = 1				
μ_1	μ_2	λ	k	h	corrected control limits h*	
0.5	4.0	0.0398	2.899	0.4306	ICARL = 500	0.3811
					ICARL = 100	0.3656
1.0	5.0	0.1354	3.2587	0.7931	ICARL = 500	0.7557
					ICARL = 100	0.7463

Table 16. The OOCARL values of the NPAEWMA chart for the six distributions. The nominal ICARL equals 500/100. Each ARL value is calculated from 10,000 simulations of run lengths. The reference sample size is m = 100 and the subgroup size is n = 1.

Shift pairs		Shift in location	Underlying distributions					
μ_1	μ_2	δ	normal	symmetric heavy tailed	symmetric light tailed	slightly skewed	slightly skewed heavy tailed	highly skewed
		in unit of 1	nominal ARL = 500, m = 100, n = 1					
0.5	4.0	0.00	494.87	493.26	499.35	500.77	498.40	501.83
		0.25	144.53	247.64	118.35	155.42	274.00	34.22
		0.50	36.16	71.94	31.01	34.54	73.01	19.35
		1.00	13.43	21.71	12.16	12.87	21.05	11.67
		1.50	8.83	13.60	8.14	8.74	13.10	8.94
		2.00	7.09	10.46	6.63	7.12	10.05	7.67
		2.50	6.33	8.85	6.11	6.36	8.61	6.96
		3.00	6.07	7.88	6.01	6.04	7.72	6.52
1.0	5.0	0.00	495.66	499.30	495.98	509.90	493.45	495.82
		0.25	186.28	314.32	154.85	238.04	352.74	27.15
		0.50	42.06	109.86	34.87	50.13	130.27	14.54
		1.00	10.76	22.20	9.46	10.89	23.14	8.03
		1.50	6.18	11.10	5.56	6.20	11.04	5.80
		2.00	4.68	7.68	4.32	4.70	7.54	4.61
		2.50	4.10	6.15	3.83	4.12	6.00	3.98
		3.00	3.78	5.29	3.48	3.96	5.16	3.58
		in unit of 1	nominal ARL = 100, m = 100, n = 1					
0.5	4.0	0.00	101.93	101.70	101.10	100.54	101.01	101.86
		0.25	49.46	68.50	44.97	53.42	72.53	18.24
		0.50	19.72	32.24	17.63	19.22	32.62	11.21
		1.00	7.98	12.62	7.31	7.70	12.41	6.79
		1.50	5.27	8.13	4.80	5.15	7.83	5.15
		2.00	4.14	6.20	3.82	4.19	5.99	4.27
		2.50	3.58	5.23	3.32	3.78	5.08	3.77
		3.00	3.24	4.67	3.09	3.43	4.52	3.43
1.0	5.0	0.00	100.70	99.49	100.16	100.27	100.49	99.42
		0.25	55.30	72.93	49.49	64.85	81.10	14.54
		0.50	21.93	37.35	18.42	23.19	42.59	8.77
		1.00	6.93	12.49	6.07	6.98	13.37	5.09
		1.50	3.97	7.00	3.56	4.00	7.04	3.75
		2.00	2.90	4.97	2.62	3.01	4.89	3.02
		2.50	2.39	3.97	2.21	2.54	3.90	2.57
		3.00	2.14	3.41	2.04	2.20	3.37	2.34

NPAEWMA chart is not directly comparable with any of these charts since the NPAEWMA chart is a Phase II chart based on a fixed Phase I reference sample, while the rest of the charts are sequential charts based on a change-point formulation. In order to make some comparisons, we chose the change-point at a relatively large number of observations, specifically at the 101th observation, to make the comparisons more reasonable among the charts. From Table 17, we can see that the two NPAEWMA control charts perform better when there is a small shift in the location parameter. Also, the NPAEWMA chart for shift pairs 1 and 5 performs well when there is a large shift to the location parameter. The parameters of the NPAEWMA charts are obtained from Table 15. The nominal ICARL values are all fixed at 500.

Example

To illustrate the proposed NPAEWMA control chart, we consider a dataset on the inside diameters of piston rings for an automotive engine produced by a forging process (Montgomery 2012). Twenty-five samples, each of size five, have been taken when the process is thought to be in-control. These are the Phase I data with m = 125. Then the process is monitored with 15 additional samples (Phase II). Table 18 shows the calculation of the NPAEWMA charting statistics and Figure 2 shows a plot of the NPAEWMA control chart for shift pair 1 and 5, with $(\lambda, k, h^*) = (0.1354, 3.2587, 0.8078)$. The $h^* = 0.8078$ is determined for a nominal ICARL of 500, when $\lambda = 0.1354$ and $k = 3.2587$, as with the AEWMA chart, when m = 125 and n = 5, using simulations. The control chart shows an out-of-control signal at the 12th test sample.

Note that the h value for the AEWMA chart was 0.7931 but the AEWMA chart assumes normality and the known parameter case. The h^* value for the NPAEWMA chart is larger, 0.8078. So the proposed nonparametric Phase II charts based on a Phase I reference sample are wider, which is expected. In order to cast a light on how sensitive the control limits are

Table 17. The OOCARL values of the NPAEWMA, NAE, HLE, McC, and HDC charts for the N(0, 1) distribution[1].

N(0, 1) vs. N(δ, 1)										
	NPAEWMA (μ_1, μ_2)			HLE Λ			McC k			
shift δ	(0.5, 4)	(1, 5)	NAE	0.05	0.1	0.2	0.05	0.1	0.2	HDC
0.25	144.5	186.3	231.0	258.0	308.0	369.0	232.0	319.0	391.0	291.0
0.5	36.2	42.1	48.9	54.4	81.8	142.0	46.8	83.5	208.0	58.7
1	13.4	10.8	13.4	13.8	12.6	13.6	14.3	12.5	17.9	12.4
1.5	8.8	6.2	7.3	9.1	7.7	6.8	9.6	7.5	6.5	6.9
2	7.1	4.7	5.2	7.3	6.0	5.1	7.8	6.0	4.7	5.1
3	6.1	3.8	4.3	6.2	5.1	4.1	6.7	5.1	3.7	4.0

[1]The OCCARL values of the NPAEWMA chart are calculated from 10,000 simulations. The OOCARL values of the NAE, HLE, McC, and HDC charts are from Liu et al. (2013).

Table 18. Calculations for the NPAEWMA control chart for the inside diameters of piston rings[a].

Phase II sample	WRS Statistics	Standardized WRS Statistics	e	Monitoring Statistics	Control limits (+/−)
1	428.5	1.2227	1.2227	0.1656	0.8078
2	346.5	0.2300	0.0645	0.1743	0.8078
3	159.5	− 2.0339	− 2.2082	− 0.1247	0.8078
4	384.5	0.6901	0.8148	− 0.0144	0.8078
5	257.5	− 0.8474	− 0.8331	− 0.1272	0.8078
6	425	1.1804	1.3075	0.0499	0.8078
7	408	0.9746	0.9247	0.1751	0.8078
8	255.5	− 0.8717	− 1.0467	0.0333	0.8078
9	485.5	1.9128	1.8795	0.2878	0.8078
10	500	2.0884	1.8005	0.5316	0.8078
11	354.5	0.3269	− 0.2047	0.5039	0.8078
12	575.5	3.0024	2.4985	0.8422*	0.8078
13	588	3.1537	2.3115	1.1552*	0.8078
14	615.5	3.4866	2.3315	1.4709*	0.8078
15	498.5	2.0702	0.5993	1.5520*	0.8078

[a]An * indicates an out-of-control signal. Average rank is used to calculate the WRS statistics when there are ties.

to the normality assumption and the estimation of parameters, corrected h* values are obtained via simulation for the NPAEWMA chart, for shift pair 1 and 5, λ = 0.1354 and k = 3.2587, and nominal ICARL = 500, when m = 50, 100, 150, 300, 500 and n = 5 for the normal distribution. Our results show that the h* values vary from 0.7931–0.8082, which indicates a reasonable robustness of the proposed NPAEWMA chart to the varying Phase I sample size.

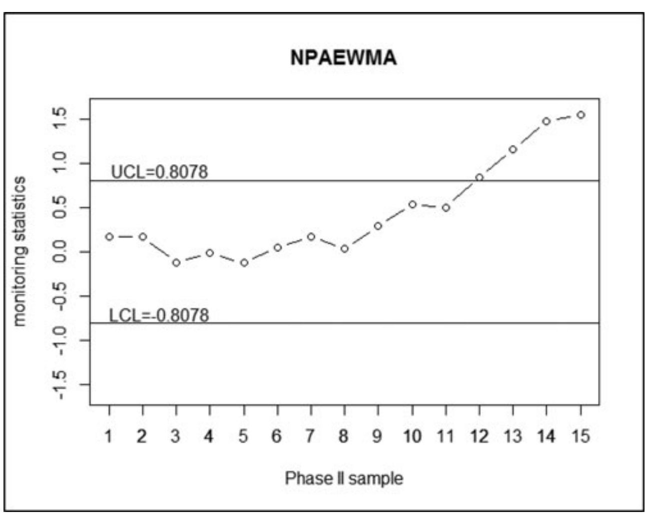

Figure 2. NPAEWMA control chart for the piston rings data.

Summary and conclusions

In this article, we examine the ICARL robustness of the AEWMA control chart proposed by Capizzi and Masarotto (2003). Our results show that as proposed, the AEWMA chart performs well under normality, but is very sensitive to the shape of the underlying distribution and hence to non-normality. Thus, if the AEWMA chart is used in situations where the normality assumption is not satisfied, there is a considerable amount of risk of getting a false alarm rate different from the advertised nominal value; it could be much higher or lower and this can have a deleterious effect on the overall monitoring regime. Motivated by this, we propose a nonparametric analog of the AEWMA chart, called the NPAEWMA chart for the unknown parameter case, based on the Wilcoxon rank sum statistic and a Phase I reference sample. The Phase II distribution-free NPAEWMA chart is shown to maintain the in-control robustness to non-normality better for distributions of different shapes and inherits the good shift detection properties of the AEWMA chart to quickly detect a shift of an unknown magnitude.

About the authors

R. Zheng is currently a Ph.D. candidate of Statistics at the University of Alabama in Tuscaloosa. She is a member of the American Statistical Association and Institute of Mathematical Statistics. Rong received her B.S. degree in Applied Mathematics from Henan University in China, in 2012 and her M.S. degree in Statistics from the University of Alabama in Tuscaloosa, in 2014. Her current research interest is application of nonparametric statistical methods to the area of statistical process control. This paper is a part of her Ph.D. dissertation work.

S. Chakraborti is Professor of Statistics at the University of Alabama in Tuscaloosa. He is a Fellow of the American Statistical Association and an Elected Member of the International Statistical Institute. Professor Chakraborti has authored and co-authored over one hundred publications in a variety of journals and outlets. He is the co-author of the highly acclaimed book *Nonparametric Statistical Inference*, 5th ed. (2010). Professor Chakraborti has contributed in a number of research areas including censored data analysis, income distribution, poverty, reliability, and general statistical inference. His current research interests include applications of nonparametric statistical methods to the area of industrial statistics and statistical process control. He has supervised over 20 Master's and Ph.D. students and has been cited for his contributions in mentoring and collaborative work with people from around the world. He has been a Fulbright Senior Scholar to South Africa and a visiting professor in several countries including Turkey, India, Holland, and Brazil. Professor Chakraborti is heavily involved in editorial work; he is currently serving his 15-year term as an Associate Editor of *Communications in Statistics* and he routinely assists a number of international journals, government and non-governmental agencies as a reviewer.

Acknowledgments

The authors are grateful to two anonymous referees for their comments which have improved the article.

References

Borror, C. M., D. C. Montgomery, and G. C. Runger. 1999. Robustness of the EWMA control chart to non-normality. *Journal of Quality Technology* 31 (3):309–316.

Capizzi, G., and G. Masarotto. 2003. An adaptive exponentially weighted moving average control chart. *Technometrics* 45 (3):199–207.

Chakraborti, S., and M. A. van de Wiel. 2008. A non-parametric control chart based on the Mann-Whitney statistic. *IMS Collections. Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen* 1:156–172.

Epprecht, E. K., L. D. Loureiro, and S. Chakraborti. 2015. Effect of the amount of phase I data on the phase II performance of S^2 and S control chart. *Journal of Quality Technology* 47 (2): 139–155.

Gibbons, J. D., and S. Chakraborti. 2010. *Nonparametric statistical inference*. Boca Raton, FL: CRC Press.

Hackl, P., and J. Ledolter. 1991. A control chart based on ranks. *Journal of Quality Technology* 23 (2):117–124.

Haynes, M. A., H. L. MacGillivray, and K. L. Mengersen. 1997. Robustness of ranking and selection rules using generalised g-and-k distributions. *Journal of Statistical Planning and Inference* 65 (1):45–66.

Hawkins, D. M., and Q. Deng. 2010. A nonparametric change-point control chart. *Journal of Quality Technology* 42 (2): 165.

Hoaglin, D. 1986. Summarizing shape numerically: the g-and-h distributions. In *Exploring data tables, trends and shapes*. ed. D. C. Hoaglin, F. Mosteller, J. W., Tukey, NY: Wiley.

Huang, W., L. Shu, and Y. Su. 2014. An accurate evaluation of adaptive exponentially weighted moving average schemes. *IIE Transactions* 46 (5):457–469.

Human, S. W., P. Kritzinger, and S. Chakraborti. 2011. Robustness of the EWMA control chart for individual observations. *Journal of Applied Statistics* 38 (10):2071–2087.

Jones, L. A. 2002. The statistical design of EWMA control charts with estimated parameters. *Journal of Quality Technology* 34 (3):277–288.

Liu, L., F. Tsung, and J. Zhang. 2014. Adaptive nonparametric CUSUM scheme for detecting unknown shifts in location. *International Journal of Production Research* 52 (6):1592–1606.

Liu, L., X. Zi, J. Zhang, and Z. Wang. 2013. A sequential rank-based nonparametric adaptive EWMA control chart. *Communications in Statistics—Simulation and Computation* 42 (4):841–859.

Mahmoud, M. A., and P. E. Maravelakis. 2010. The performance of the MEWMA control chart when parameters are estimated. *Communications in Statistics—Simulation and Computation* 39 (9):1803–1817.

Mahmoud, M. A., and A. R. Zahran. 2010. A multivariate adaptive exponentially weighted moving average control chart. *Communications in Statistics—Theory and Methods* 39 (4):606–625.

McDonald, D. 1990. A CUSUM procedure based on sequential ranks. *Naval Research Logistics* 37 (5):627–646.

Montgomery, D. C. 2012. *Introduction to statistical process control*. New York, NY: John Wily& Sons.

Rayner, G. 1999. Statistical methodologies for quantile-based distributional families. Ph.D. Thesis, Queensland University of Technology (QUT).

Reynolds Jr, M. R., and Z. G. Stoumbos. 2006. Comparisons of some exponentially weighted moving average control charts for monitoring the process mean and variance. *Technometrics* 48 (4):550–567.

Saleh, N. A., M. A. Mahmoud, and A. S. G. Abdel-Salam. 2013. The performance of the adaptive exponentially weighted moving average control chart with estimated parameters. *Quality and Reliability Engineering International* 29 (4):595–606.

Saleh, N. A., M. A. Mahmoud, L. A. Jones-Farmer, I. Zwetsloot, and W. H. Woodall. 2015. Another look at the EWMA control chart with estimated parameters. *Journal of Quality Technology* 47 (4):363–382.

Shu, L. 2008. An adaptive exponentially weighted moving average control chart for monitoring process variances. *Journal of Statistical Computation and Simulation* 78 (4):367–384.

Simões, B. F., E. K. Epprecht, and A. F. Costa. 2010. Performance comparisons of EWMA control chart schemes. *Quality Technology & Quantitative Management* 7 (3):249–261.

Tang, D., K. Chen, X. Chen, H. Liu, and X. Li. 2014. Adaptive EWMA Method based on abnormal network traffic for LDoS attacks. *Mathematical Problems in Engineering* 2014: 496376. DOI: 10.1155/2014/496376.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* 1 (6):80–83.

Woodall, W. H., and M. A. Mahmoud. 2005. The inertial properties of quality control charts. *Technometrics* 47 (4):425–436.